DOCUMENT RESUME

ED 040 385 AL 002 436

AUTHOR Streiff, Paul R.
TITLE Some Criteria for Designing Evaluation of TESOL Programs.
PUB DATE Mar 70
NOTE 20p.; Paper given at the fourth annual TESOL Convention, San Francisco, California, March 18-21, 1970

EDRS PRICE EDRS Price MF-$0.25 HC-$1.10
DESCRIPTORS *English (Second Language), Evaluation Criteria, *Evaluation Methods, Evaluation Needs, Objectives, Performance Specifications, *Program Evaluation, Standards

ABSTRACT

The author discusses the growing body of information in which can be found criteria for applying evaluation procedures to programs in TESOL (teaching English to speakers of other languages). One problem is that often educational goals are not stated explicitly or specifically; implicitly stated or implied goals are difficult to communicate to others or to analyze. Evaluation goals should be distinguished from the roles evaluation may play, which may be as part of (1) teaching training activity; (2) curriculum development; (3) field experiments in learning theories; (4) selection of materials; and (5) reward or punishment; e.g., in an executive training program. Evaluation may also have a role in on-going curriculum improvement raising such questions as "Is the curriculum at this point getting across the intended instructional objectives?" "Is it taking too large a proportion of the available time?" Evaluation needs in TESOL are extremely varied, in terms of target population and objectives and the roles evaluation is asked to assume within situations. There is a need for people within the TESOL Association to become better informed in matters of evaluation. A Center of Evaluation of TESOL Programs should be established by the TESOL Association as a "matter of urgency and self-interest." (AMM)

PRINTED BY THE STANDARD REGISTER COMPANY, U. S. A.
FORM 0801

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED040385

SOME CRITERIA FOR DESIGNING EVALUATION OF TESOL PROGRAMS.

Paper presented by Paul R. Streiff, 1970 TESOL CONVENTION

INTRODUCTION:

The title of my talk here today may be somewhat misleading. It might be taken to mean that I am going to give you the master design for evaluating any and all TESOL projects. I assure you this is not the case. It isn't possible for one design to meet the very diverse evaluation needs of TESOL programs everywhere, except in the very broadest and most general sense. Evaluation is a complex process, and as a discipline in its own right has only recently begun to emerge as a field apart from educational research. New concepts, procedures, and instruments of evaluation are evolving to meet new needs and conditions, but there aren't enough trained evaluators yet to meet the increasing demands for services, nor are there enough ready-made models to meet specific evaluation needs.

There is, however, a growing body of information in which we can find criteria for applying evaluation procedures to TESOL programs. I will enumerate some of these criteria as one objective in this talk.

Another objective derives from recent events which brought this organization directly into confrontation with matters of evaluation concerning TESOL programs. A short time ago the Bureau of Indian Affairs asked TESOL for assistance in evaluating ESL instruction in Navajo Area schools following several years of intensive efforts to make ESL an integral part of the curriculum at all levels. I have no information on the results of those evaluation efforts, but the important factor here is that ESL has for several years been viewed as essential to the improvement of instruction for many segments of the American school population. Teachers and administrators who have been involved in its implementation, as well as those who are consider-

AL 002 436

ing its use in their schools, are asking for evaluation of its effectiveness. It is logical to assume that they as well as others will look to the TESOL organization for help in finding answers, since we became an organization to provide leadership and support for the many and divers attempts to improve English language competence and performance of non-English speakers. As a second objective of this presentation I will recommend a course of action to be taken by TESOL whereby we can assist our membership in meeting increasing demands for effective evaluation of their contributions.

PART I

EVALUATION: A BRIEF OVERVIEW.

Evaluation in education has recently come into sharp focus in many contexts and instructional situations. There is a great deal of thrashing around and agonizing right now, especially prompted by evaluation requirements in federally funded programs. There is much demand that more effective evaluation be performed. Those who are out there on the firing line trying to interpret what a legitimate evaluation design should look like are, in many cases, quite at a loss as to where to turn for help. How does one develop a design and translate that design into data that may effect whether or not a given program receives continuation funding? How does one determine whether this or that set of instructional materials should be purchased? Whether the expenditure of funds for a specific consultant or firm contracted to design a teacher training workshop is justified?

Demands for evaluation in education have been with us for as long as there has been such a thing as formal education, I'm sure. People have always wanted to know something about the outcomes of their efforts to pass on all those good things about the society into which their children are born, and to pass it on in ways that are acceptable to 'them', 'them' meaning the ones who pay the bills. For the most part, in the past, however, evaluation has been seen as necessary in deciding who, among the learners, will be rewarded and who will be punished----who will be allowed to go on in the system and who will have to be cut from the roster. In this view of evaluation it is really synonymous with measurement, and often the two terms have been used interchangeably. Measurement, however, should be used to refer to

quantitative descriptions of behavior, things, or events. It may be defined as descriptions of behavior that can be expressed in numbers, while evaluation has a broader scope, which includes measurement. Ralph Tyler put it this way, enumerating specific criteria:

"The purpose of educational evaluation is expository: To acquaint the audience with the workings of certain educators and their learners. It differs from educational research in its orientation to a specific program rather than to variables common to many programs. A FULL evaluation results in a story, supported perhaps by statistics and profiles. It tells what happened. It reveals perceptions and judgements that different groups and individuals hold--obtained, I hope, by objective means. It tells of merit and shortcomings. As a bonus, it may offer generalizations for the guidance of subsequent educational programs. Evaluation requires the collection, processing, and interpretation of data pertaining to an educational program.....(and) two main kinds of data are collected: 1) objective <u>descriptions</u> of goals, environments, personnel, methods and content, and 2) personal <u>judgements</u> as to the quality and appropriateness of those goals, environments, personnel, methods, and content.(1)

In the paper presented by Mr. Wooley we heard about an evaluation model developed by the Center for the Study of Evaluation at UCLA. The definition of evaluation offered by Dr. Marvin Alkin, Director of the Center, reflects the major aspects of evaluation which, I believe, would be agreed upon by a majority of authorities in the field. Alkin defines evaluation as:

> The process of ascertaining the decision areas of concern, selecting appropriate information, and collecting and analyzing information in order to report summary data useful to decision-makers in selecting among alternatives.(2)

Probably the most significant single factor is the current emphasis on evaluation was the passage of the Elementary and Secondary Education Act of 1965. Titles I, III, and VII of that act all contain requirements that local districts receiving such funds must evaluate the effectiveness of the educational efforts for which grants are made. Many schools in a majority of states indicated at the outset that they had no means readily available, according to Ralph Tyler, for conducting such evaluative studies.(3) Review of projects for continuation funding has emphasized the tremendous lack of personnel and facilities nationwide to meet those requirements: Requirements based ultimately on the fact that the source of dollars for funding innovative programs, the United States Congress representing the American taxpayer, is insisting on straightforward answers to questions like, "What did we get for the money?" This lack of available means for conducting necessary evaluative studies has led to the establishment of several centers that are developing new theories and procedures. The center at UCLA is one of them, and Mr. Wooley's appearance on this program is quite significant for TESOL. He is presently involved, as a representative of the Center for the Study of Evaluation, in matters of evaluation involving programs funded under Title VII of ESEA, the Bilingual Education Act. In spite of some opinion to the contrary, ESL must be a very important part of any bilingual education program in American Schools: A bilingual program elsewhere may involve any combination of languages imaginable for any number of legitimate reasons, but the language of Title VII specifies that English must be one of the languages of instruction in all programs. Mr. Wooley and other evaluators are therefore going to be involved in learning something about TESOL while providing expertise and services in matters of evaluation. We are aware, as a result of his

presentation and presence here, that:

a) There is a considerable body of information available on evaluation.

b) There are some Centers already established for the study of evaluation.

c) Models for evaluation have been developed and can serve in the development of other models.

In regard to the last point I will briefly mention two other models you may wish to investigate. One is called the Pittsburgh Discrepancy Model, developed by Robert Provus and reported in the 1969 Yearbook of the National Society for the Study of Education.(4) The other was developed by Professor W. James Popham and is entitled PROGRAM FAIR EVALUATION-SUMMATIVE APPRAISAL OF INSTRUCTIONAL SEQUENCES WITH DISSIMILAR OBJECTIVES. The Popham model is reported in a paper available from the Southwest Regional Laboratory for Educational Research and Development of Los Angeles.(5)

What is the significance of all this for TESOL programs and the TESOL organization? We have only to briefly review the short history of this organization to realize that it has a great deal to do with whether we are honored for our contributions or become one of those overworked but poorly understood sets of labels that have too quickly become meaningless to many and even anathema to others. Along with all aspects of curriculum improvement effort we are in the position of having to justify our claims on empirical bases, and to do this we need to understand what is meant by evaluation and come up with specific plans for designing appropriate evaluation models which will effectively establish the importance of TESOL efforts in terms of measurable learner behavior. It is gratifying to know that evaluators are concerning themselves with ESL matters

in connection with Title VII programs, but those programs represent only a small part of the concerns reflected by our membership, and evaluation demands are upon us all. Present facilities and personnel are not adequate to meet our needs.

EVALUATION NEEDS IN TESOL

One point on which there seems to be general agreement among evaluators is that the evaluation process must begin with a determination of what is to be evaluated. Though this seems almost too simplistic to mention, the fact is that all too often there is little or no specification by those determining the design and content of curriculum of what the outcomes of instruction are intended to be. To be sure, it hardly seems possible that anyone could be involved in the teaching business for any period of time without having set some specific goals by which his actions are guided, yet those specifications have, for the most part, been implicit in one's actions rather than stated in precise and explicit form. Benjamin Bloom, in a recent paper had this to say:

> "Unfortunately, specifications which are implicit are difficult to communicate to others, they are rarely analyzed and clearly revised, and they do not serve as clear guides to particular decisions or actions. Implicit specifications may shift without the educational worker being clearly aware of any change, and, because of poor communications, the attainment of the specifications may defy any attempt at systematic appraisal.(6)

And most important for the TESOL effort, and for this organization, he said,

> "Trust in professionals is a highly desirable goal for any field, including education, but each profession must either police itself, if it is to merit the confidence of the public that supports or uses it, or expose itself to external scrutiny when the confidence of the public is impaired."(7)

If the purposes and specifications for a curricular area like ESL are not explicit, then it is possible for them to be altered by social pressures, by new fads and fashions, and by new schemes and devices which may come and go with momentary shifts on the educational scene. Implicit purposes are difficult to defend and impossible to evaluate.

But what does this have to do with TESOL programs in general? Or TESOL programs in particular? Is it possible that those who have been responsible for the design and implementation of ESL programs in American schools don't know about this business of explicit specification of learner post-instruction behavior? I'm sure we could find evidence that this is true as well as evidence to the contrary. Where such explicitness is lacking the first and most important task we have, in order to make possible the evaluation of TESOL Programs, is to see to it that all of our professionals acquire the skills necessary to generate appropriate specific objectives. Once we have managed to accomplish this task at least the foundation for evaluation has been laid and we can proceed from there.

But is this likely to be the most efficient course to follow in attempting to assist those of our number who need help in adequately assessing the impact of their efforts? If we were to take a careful look at what our membership does in education I'm sure it would come as no surprise to most of us that interests and involvements range from pre-school through adult education, basic and otherwise; from teaching to monolingual speakers of another language to the teaching of students who have managed, unfortunately, to mix two language systems, and as a result perform with difficulty outside their family and community in either language: from teaching English as a new language to teaching standard English as a second dialect. We have such a range of interests, in fact, that it would be most difficult for anyone to cover all areas adequately with even the most comprehensive set of language learning objectives, and indeed these would more than likely prove to be superfluous to the evaluation needs of any given program.

Looking still more closely at the all important question, "What is to be evaluated?" we find that within each level and interest area we encounter a considerable assortment of roles to be played by evaluation to meet an increasingly complex assortment of demands. In one of the important papers on evaluation in the past several years Michael Scriven made a point of distinguishing between goals of evaluation as against roles. He stated that "goal" activity consists simply of the gathering and combining of performance data with a weighted set of goal scales to yield either comparative or numerical ratings. These data are used in the justification of a) the selection of objectives, b) the data gathering instruments, and c) the weightings.

The ROLE of evaluation, however, varies tremendously between particular educational contexts. It plays a role as part of a teacher training activity; as part of the process of curriculum development; as an essential element in the conduct of a field experiment connected with the improvement of learning theory; as part of an investigation preliminary to a decision about the purchase or rejection of materials; as a preliminary activity in the reward or punishment of people as in an executive training program.

According to Scriven, evaluation can and usually should play several roles. Not only can it have several roles with respect to one educational enterprise, but with respect to each of these it may have several specific goals. It may have a role in the on-going improvement of the curriculum, and with respect to this role several types of questions or goals may be raised, such as; is the curriculum at this point really getting across the intended instructional objectives. Is it taking too large a proportion of the available time to make its point? etc.(8)

Professor Popham's design, mentioned earlier, reaffirms an essential aspect of evaluation that distinguishes it from experimental research and which TESOL personnel must keep in mind: That is, that goals of evaluation always include the estimation of merit, worth, value, etc. In calling for the assignment of values to objectives Popham says, "This, of course, is sticky business, and it would be pleasant to avoid the imprecision and subjectivity of making value judgements altogether. Unfortunately, we can't. At this point (in the process) the evaluator must reach a position regarding the respective worth of the ....sets of objectives. Ideally, the relative value of the sets could even be quantified."(9)

In briefly summarizing this section on evaluation needs in TESOL programs we may say that:

1) Needs are extremely varied, in terms of target population and objectives.
2) Objectives are in need of careful specification in each situation. What will the learner be able to do at the end of a particular lesson that he was unable to do before?
3) Needs are extremely varied in terms of the roles evaluation is asked to assume within situations.
4) There is a need for people within the TESOL Organization to become better informed in matters of evaluation.

At this point there are certain clarifications which ought to be made about the process of evaluation, and several criteria established for any involvement with that process. These criteria, if incorporated into the TESOL context, should make it possible for this organization to make important contributions to the systematic improvement of educational programs.

a). First, the evaluator need not necessarily have participated in the planning of a program in order to be effective. Many programs have not been planned, at least not in terms of careful specification of outcomes as learner-post-instruction behavior. If these programs are to be evaluated a strategy must be found which doesn't depend on participation during the planning stage. There are such designs, and others can assuredly be built.

b) Second, it isn't necessary to wait three to five years before an evaluative judgement can be passed on a new program. A whole series of enroute judgements must be made to decide if one should review a stage, go on to the next, or terminate the project.

c) Third, it is not inevitable that conflict exists between evaluator and program people. Both have the same mission: either to continue and improve a program or to reject it as soon as reliable evidence is gathered that its success potential is minimal or at least very low.

d) Evaluation activity should not be viewed as getting in the way of program activity since it is a necessary part of program development and should therefore be thought of as complementary.

e) Good evaluation does not depend on the adoption at the outset of a sound <u>experimental</u> design for the program being evaluated.. Experimental design is irrelevant to evaluation until a program is in its final stages of development, and if the evaluation was properly conceived and conducted the project may have been legitimately terminated long before reaching that point. In order to improve a program in its formative stages there must be opportunity to improve it on the basis of information gathered through experience and judgements made on the basis of that information.

> A formal experimental design would deny the opportunity to make changes while a program is in its dynamic stages of growth.(10)

Sound evaluation practice provides administrators and staff with information they need and freedom to act on that information. (Put simply, evaluation is the process of a) agreeing upon program standards. b) determining whether a discrepancy exists between actual outcomes and the agreed upon standards, and c) using that information to identify weaknesses and recommend next steps based on d) data-based value judgements.) The process of evaluation being complementary, instructional decision-making can be viewed as a series of steps in which the following activities take place:

1. Objectives are selected. Objectives guide choices about processes and are the bases used in judging final outcomes. In TESOL programs this means specification of terminal behavior which in some way reflects attainment of very specific linguistic objectives.
2. Instructional sequences are selected which are to produce the learning outcomes specified in the objectives: That is, strategies for instruction are selected on which learning can be predicted.
3. Learning outcomes reflected in learner post-instruction behavior are measured and/or described.
4. Outcomes are compared with specifications called for initially in the objectives.
5. If the outcomes attained fall short of those specified by the objectives, alternative decisions are called for.

While the first two steps are usually up to the decision-maker within the system, it is the responsibility of competent evaluation effort to objectively carry out steps three, four,and five.

A PROPER CONCERN OF TESOL

I have spent considerable time setting out some criteria for evaluation of TESOL programs. I now suggest that evaluation, and the provision for assistance to its members in evaluation matters are or should be proper concerns of this organization for the following reasons:

a) We have a large membership with only loosely similar objectives: They have something to do with teaching English to non-English speakers.

b) Specification of objectives in behavioral terms for TESOL programs is not a simple matter and since it is essential to evaluation procedures we might begin here. This is a difficult task for any given program, and assistance is necessary and long overdue.

c) So far TESOL programs have mostly operated on the sure feeling that, like most education efforts, we just KNOW it's good and right and making a significant contribution. Gut feelings aren't good enough any more. We're going to have to tighten up.

d) Yet there is probably a great deal of useful data, carefully recorded but unused because it has never found a home where it could be utilized for evaluative purposes. The individual struggling at the local level needs a strong assist from the outside many times, in order that his efforts in ESL aren't eclipsed by newly popular concerns.

e) It has been said that the effect of the Elementary and Secondary Education Act may very likely be that establishing evaluation as a necessary building block of American educational reform may eventually prove to have greater impact on education than the programs themselves. We may find that, through effective evaluation TESOL programs may look very different than we've known them. Our assumptions about TESOL programs may lead to new avenues of approach, method and technique.

WHO IS GOING TO EVALUATE TESOL PROGRAMS?

I have already commented on the involvement of the Center for the Study of Evaluation at UCLA in the evaluation of Title VII projects, and since these projects by definition must have TESOL components we realize that evaluators are becoming interested in TESOL. However, few graduate programs provide training for evaluation specialists. Evaluation specialists are and will be increasingly needed throughout education, but for now there just aren't enough bodies to go round. Rather than waiting for the supply to catch up we might better devise a plan for systematically developing and providing evaluation expertise and needed assistance from within the TESOL organization. For the immediate future there is much that can be done by administrators, teachers, and anyone connected with TESOL programs to help improve the quality of TESOL through effective evaluation procedures.

CONCLUSIONS:

On the national level I suggest that the TESOL organization jump right into the matter of evaluation as a matter of urgency and self-interest. WE need to know how to find out how we are doing because others are wanting to know---and by others I mean those who are paying the bills.

I propose that it is possible for this organization to obtain the funding necessary to establish a Center for the study of Evaluation of TESOL and TESOL related programs. The present administration has made it very clear that funding will be made available to undertake more relevant Research and Development activities; to find out in systematic fashion what does and doesn't work. This would be the focus of such a center. Our concerns span a tremendous range of the national educational interest. We are a new and dynamic organization, and establishment of such a center would make possible an increasingly significant contribution to education.

The Center for Evaluation of TESOL Programs should be set up to serve functions like the following:

a) Establish a research and development focus for TESOL.

b) Seek out hard data from TESOL projects and programs where already available.

c) Respond to requests from on-going projects for help in designing evaluation models for part or all of the project, and assist in implementation.

d) Assist in designing legitimate 'new' studies involving bilingualism and bilingual education particularly. There is no other organization as logically equipped to help settle some of the persistent questions.

I might suggest as a starter the replication of Wallace Lambert's recent exciting study into some cognitive consequences of following the curricula of grades one and two in a second language.

e) Provide for dissemination of findings, and particularly of evaluation designs.

In conclusion, we may summarize as follows:

1) Evaluation is a complex process, but very much the order of the day throughout education.

2) TESOL programs are being asked to evaluate their efforts. Those programs are very diverse in nature and focus.

3) While there is considerable activity in the field of evaluation, and some general criteria for designing evaluation are available, there is a great shortage of personnel and facilities for getting the job done.

4) The TESOL organization should establish a Center for Evaluation of TESOL programs to provide assistance to its membership for systematically improving their projects.

FOOTNOTES

1. Ralph W. Tyler, Editor, NSSE Yearbook Part II 1969. Educational Evaluation: New Roles, New Means. University of Chicago Press, Chicago, 1969.

2. Marvin C. Alkin. Towards an Evaluation Model: A Systems Approach. UCLA Center for the Study of Evaluation of Instructional Programs (CSEIP). Working Paper No. 4, February, 1968.

3. Tyler. op. cit. pp. 1

4. Robert Provus. NSSE Yearbook Part II 1969. Evaluation of Ongoing Programs. University of Chicago Press, Chicago, 1969. Chapter XI, pp. 242.

5. W. James Popham. Program Fair Evaluation-Summative Appraisal of Instructional Sequences with Dissimilar Objectives. Research Memorandum SWRL (Southwest Regional Laboratory) September, 1968.

6. Benjamin S. Bloom. NSSE Yearbook (as above) Some Theoretical Issues Relating to Educational Evaluation. Chapter III, pp. 26, etc.

7. Ibid.

8. B. Othanel Smith, Editor. See especially: Michael Scriven. The Methodology of Evaluation. AERA Monograph Series on Curriculum Evaluation. Perspectives of Curriculum Evaluation. Rand McNally & Co. 1969. pp. 40-43.

9. Popham. op. cit.

10. Provus. op. cit. pp. 282.

TESOL 1970 Session: Evaluation of ESL Programs
Saturday, March 21, 2:15